Representing the Effect of Multiple Alternatives and Information Strength on Confidence in Perceptual Decision Tasks

A Thesis

Presented to

The Academic Faculty

by

Brooklyn Crabtree

In Partial Fulfillment

of the Requirements for the Degree

B.S. Neuroscience with Research Option in the

College of Sciences

Georgia Institute of Technology

December 2020

Representing the Effect of Multiple Alternatives and Information Strength on Confidence in Perceptual Decision Tasks

Approved by:

Dr. Dobromir Rahnev, Advisor
School of Psychology
Georgia Institute of Technology

Signature: DocuSigned by: Dobromir Rahnev
57F8B5A402CE45A...

Date Approved: 12/7/2020

Dr. Mary Holder
School of Psychology
Georgia Institute of Technology

Signature: DocuSigned by: Mary Holder
A42D14CA117F4AB...

Date Approved: 12/7/2020

Contents

Abstract

Confidence in perceptual decisions is a baseline for quantitatively measuring metacognitive processes in psychology. Most researchers limit the stimulus to two choices, assuming that the mental process summarizes the likely accuracy of all choices to determine confidence in the decision. The purpose of this study was to determine whether multiple alternative choices, with varying levels of information strength for each choice, follow the same mental statistics as similar two-alternative forced choice (2AFC) tasks. If the differences between information strengths for each of the multiple choices had a direct effect on confidence, then presenting higher and lower differences of information strength between the correct choice and the incorrect choice would result in corresponding higher and lower confidence ratings. Participants were shown multicolor clouds of dots made up of three colors, with one dot color (dominant) being more abundant than the others. Participants decided which color was the dominant color for each cloud, then indicated their confidence in that decision. The overall information strength, dominant-secondary strength difference, and dominant-tertiary strength difference all had significant main effects on both confidence and accuracy. The overall strength had the largest effect size for confidence, with more information strength resulting in higher confidence ratings. The dominant-secondary strength difference had the largest effect size for accuracy, with a larger difference between dominant and secondary color strengths resulting in higher accuracy rates. Further investigation on how the brain defines relevant stimuli in an environment and processing of multiple choices must be conducted before developing computational models for confidence.

Keywords: metacognition, decision making, confidence, perceptual decisions, 2AFC, uncertainty

Introduction

Perceptual decision making is the process of using sensory information from the environment to choose and act upon one option from a list of possible alternative options. Confidence in these perceptual decisions involves the self-evaluation of that decision's quality to redirect future choices to be of equal or higher quality. This confidence in perceptual decision making drives vital interactions in society, but lack of understanding of the mechanism for these processes leads to severe negative consequences. For example, business stock analysis requires high-stakes perceptual decisions to be made regularly. In large businesses, complex perceptual decisions that involve interpreting large fluctuations in stock investments often lead to employees expecting a change too extreme in value because of the high volatility (Zylberberg et al., 2016). When professional stock analysts overreact in predicting poor earnings forecasts, stock prices often drop dramatically and result in severe investment losses (De Bondt & Thaler, 1990). While the role of overreaction to highly volatile stimuli is understood, the exact mechanism of how sensory information modulates the confidence process is not well developed in psychological literature.

Computational models of the confidence process in decision making need further development to increase external validity and be applicable to perceptual scenarios beyond the laboratory environment. The current literature is limited by study designs varying the strength of sensory information over only two possible choices. In this design, participants are presented with a perceptual task where a stimulus is more accurately one choice over another, like more left than right or more blue than red (Samaha et al., 2016; Boldt et al., 2017). The participants decide which choice was correct, and then rate their confidence in that decision. Most researchers in the field argue that the best basis for modeling this confidence data is in Bayesian posterior probability distributions (Zylberberg et al., 2014; Sanders et al., 2016; Meyniel et al., 2015). The Bayesian basis suggests that human confidence is processed by computing the joint probability of potential decisions within a given environment. With an additional shift accounting for human error, this mental computation would follow the statistical computation exactly.

However, the given environments for these studies pose a modeling issue. The researchers proposing this confidence model limit the decision choices to two alternatives and have not explored what occurs when more alternatives are added (Zylberberg et al., 2014; Sanders et al., 2016; Meyniel et al., 2015). According to a cognitive study by Li and Ma, this Bayesian distribution of confidence only applies to the two-choice decision making tasks, and adding a third alternative for decision making disrupts this distribution so that it does not model confidence data (2020). They suggested that confidence was not based on the holistic probability of an event, as suggested by Bayesian model researchers, nor was it based on the individual probability of any alternative choice. Instead, confidence would theoretically be represented by the differential probability of an event, where the difference between the two highest choices is used to determine confidence. As reality often presents stimuli in multiple alternatives for humans to make perceptual decisions on, this differential probability model must be tested for confidence in perceptual decision making tasks.

In this study, we will present multiple-alternative perceptual decision making tasks to participants and explore components for future computational models to understand how the confidence process works with additional alternative choices introduced. We predict modeling confidence in perceptual decision making with increasingly complex behavioral scenarios would increase our knowledge of human confidence in terms that can be actively applied to both stronger neuroimaging models of confidence as well as stronger training programs that increase quality decisions in high-stakes perceptual scenarios.

Literature Review

Differences Between Confidence and Perceptual Decision Making

Confidence about decisions involves subjective evaluation of the quality of the decisions to redirect future decisions (Boldt et al., 2017). The difficulty in defining confidence quantitatively comes from analyzing a subjective evaluation accurately, distinguishing the metacognition of this evaluation from the quality of the decision itself, and ensuring the experiment can be applied to the real world with high external validity (Zylberberg et al., 2016; Spence et al., 2018; Boldt et al., 2017).

Researchers in the field typically look to address at least one of these issues in each research study. The accuracy of confidence analysis tends to be addressed by creating a computational model that collects stimulus and decision inputs to correctly predict the resulting confidence rating from the participants (Kiani et al., 2014; Li & Ma, 2020; Zylberberg et al., 2016). The difference between confidence and the decision quality itself has often been distinguished through varying the amount of perceptual information presented to participants and determining the difference between confidence levels and decision accuracy (Desender et al., 2018; Spence et al., 2018; Vlassova et al., 2014).

Having a high external validity within the experimental design, however, is not often considered in most studies. Many researchers use the common two-alternative forced-choice (2AFC) tasks that simplify decisions to distinguish between two choices, such as yes-no or left-right (Boldt et al., 2017; Desender et al., 2018; Sanders et al., 2016). Few researchers have questioned whether presenting additional alternatives would invalidate the models created from the results of 2AFC tasks when applying the model to the real world (Li & Ma, 2020), but these questions must be considered to effectively evaluate the metacognitive process of confidence in order both to determine the best methods of increasing quality decision making and human confidence as well as to discover the neural pathways responsible for confidence.

Confidence Modeling: Bayesian Posterior Probability Distribution

The most commonly referenced model in the literature for representing confidence is the Bayesian posterior probability distribution. Zylberberg, Roelfsema, and Sigman used a Bayesian model to display the difference between inaccurate perception and confidence judgments by comparing the participant confidence reports to the standard error as dependent on internal evidence variance (2014). The model depends on differences in stimulus variance to determine differences in levels of confidence; however, the variance in the design was set up in categorical levels of low, medium, and high rather than a continuous scale of 0% to 100%. As the model only uses these three levels, it is impossible to determine whether any more complex relationships exist between stimulus variance and confidence at more ranges of variance and significantly limits the utility of this computational model.

Unfortunately, most studies referencing Bayesian distributions as their model do not define their parameters even as specifically as Zylberberg and colleagues did in their 2014 study. Several studies just refer to the distribution as the mathematical approximation of confidence in human decision making without fitting the model to any data to ensure a good model fit (Boldt et al., 2017; Meyniel et al., 2015; Sanders et al., 2016). These studies do not add a quantitative understanding of the confidence process but only continue to suggest a trend without any verification of computation. In short, most studies describing confidence as a Bayesian distribution do not create a testable and precise computational model for analyzing the metacognitive process of confidence.

Information Strength

While distinguishing between the decision making and confidence processes presents several design challenges for researchers, many studies found success in controlling for the strength of stimulus information used in the decision making process, like the degree of angle left or right. Spence, Mattingley, and Dux were able to find distinguishing relationships for decision accuracy and confidence, where accuracy did not change significantly over informational strength while confidence significantly decreased with decreasing informational strength (2018). In this study, two stimuli were presented, and researchers asked questions about the brightness or direction after viewing both stimuli. By increasing the direction variability, and therefore decreasing information strength about direction, researchers observed that participants rated lower confidence in brightness decisions without decreasing accuracy of brightness decisions.

In another study, Zylberberg, Fetsch, and Shadlen found that increasing noise, or increasing information strength of the irrelevant stimuli, would decrease accuracy to a small effect and increase confidence to larger effects in both human and monkey decision making (2016). Zylberberg’s 2016 study falls under the same criticism as his 2014 study in that the noise, the dependent variable upon which the conclusions depend, is reduced to an unnaturally categorical variable with the levels “low” and “high” rather than a continuous variable with a wider range of observational power, suggesting the effects are weaker and should be reevaluated with a more precise experimental design. Limitations aside, both studies' data suggest that there is a direct positive relationship between irrelevant stimuli and confidence, but do not suggest a relationship between relevant stimuli and confidence and can only prescribe a relationship between the ratio of relevant to irrelevant stimuli and confidence due to their experimental designs as 2AFC tasks in a two-dimensional design.

Two-Alternative Forced Choice Task

The standard experimental design for perceptual decision making and confidence studies follows the 2AFC task design where participants make a decision between two choices from the perceptual stimulus and then rate their confidence about the decision made. In the Boldt, de Gardelle, and Yeung study in 2017, participants chose if the summary of colors presented had more red or more blue. In the Desender, Boldt, and Yeung study in 2018, participants followed the same procedures as the Boldt 2017 study by choosing between the colors being red or blue in summation, with the additional option of requesting further information in some trials. In the Meyniel, Sigman, and Mainen article, the authors describe the traditional design involved with motion perception decision making being summarized into motion vectors going either left or right, with varying degrees to the left or right (2015). Sanders, Hangya, and Kepecs presented participants with an audible stimulus in both the left and right ear and had participants choose on which side the stimulus was playing faster (2016). Even in Spence, Dux, and Arnold's 2016 study, which involved 10 direction conditions with 4 ranges of variance, the participant choice was still a summation of whether the direction was moving towards the left or the right of center. Most studies introduce complexity into the task with variance in information strength and conditions, but participants ultimately only have two options to choose between.

Li and Ma critiqued this design in their 2020 article by creating a study where three alternative choices are presented in response to their cognitive stimulus in order to develop a confidence model from three different options rather than the typical two options. In their model, they suggest that confidence is not determined by the participant making the most correct choice out of all options, but by participants subjectively making the best choice out of the two options they see as most likely. The difference-based model, based on the best two choices, better fit the data than a corresponding Bayesian distribution model, based on all presented choices. While Li and Ma's study used a cognitive task instead of a perceptual task, their conclusion has strong implications for the 2AFC dominated perceptual decision field. As no perceptual studies have addressed this question, research must be done to fill this gap in scientific knowledge, expand the understanding of the process of confidence, and validate the application of this understanding to the external world.

Current Study

The current study will address the limitations in confidence modeling, information strength, and multi-choice experimental designs by investigating components of confidence in perceptual decision-making. Using a three-alternative forced choice design with twelve conditions of information strength, this study will collect and represent the data that future computational models of confidence should fit. This study will provide a more conclusive perspective of the effects that multiple-alternatives and varying information strength have on confidence in perceptual decision-making tasks.

Methods

Participants

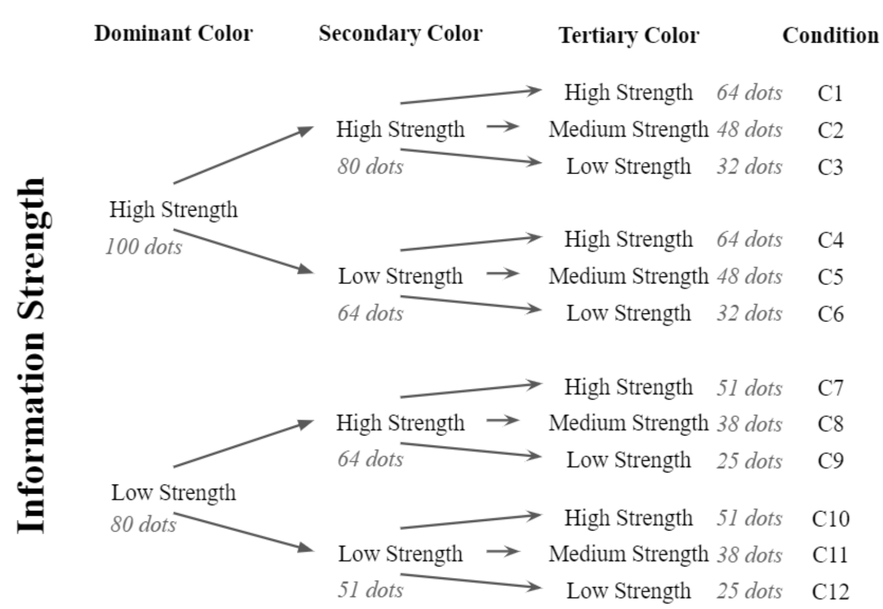
Students from the Georgia Institute of Technology (n = 15) between the ages of 18 and 40 were recruited to participate in this study at the Center for Advanced Brain Imaging via the Georgia Tech SONA psychology experiment scheduling system. The exclusionary criteria for the study required participants to be able to distinguish between different colors and complete both sessions of the study; there were three participants who completed only one session, so their data were excluded from this study. The sample was made up of 15 Georgia Tech students (8 females) aged 18 to 21 years old with an average age of 19.3 with normal (n = 11) or corrected-to-normal (n = 4) vision. Consent was approved according to guidelines from the Georgia Tech Institutional Review Board, and the experiment began after participants completed the consent forms. For their participation, students received one SONA course credit per hour for a total of two course credits.

Design

This experimental study manipulated the information strength of three visual stimuli in a decision-making task to determine the effect on perceptual accuracy and confidence. There were three levels of color: dominant, secondary, and tertiary. Dominant had the greatest abundance of dots on-screen, while secondary had fewer dots and tertiary had the fewest number of dots. There were twelve conditions based on information strength for each color level with manipulations on the difference in strength between each level, as demonstrated in Figure 1 below.

Figure 1.

Conditions for Information Strength Differences at Each Color Level



Note. C1-C12 demonstrates the twelve conditions of the experiment and their corresponding information strength of the dominant, secondary, and tertiary color. This information strength is specifically represented by the number of dots of that color on screen. For example, C5 would have 100 dots of the dominant color, 64 dots of the secondary color, and 48 dots of the tertiary color. These differences between color levels should demonstrate the relationship of three stimulus choices and their information strengths on confidence in perceptual decision-making tasks. It should be noted that the first six conditions (C1-C6) have more dots overall than the last six conditions (C7-C12). This is designed to distinguish the effects of overall information strength presented in a stimulus from the effects of color level information strength in a task.

The information strength was defined by the number of dots. Overall, this was represented by the number of dots visible on the screen during a given period, so high overall information strength would have more dots present like in condition C1. This strength was also represented for each color level, where high level information strength in a level had more dots present than a condition with a low-level information strength on the same level, like conditions C2 and C5.

Accuracy was defined by whether the participant's response was correct or incorrect for each trial. Confidence was defined by a Likert scale response of 1 to 4, with 1 being low and 4 being high. The study is a within-groups design, so each participant is presented with all twelve conditions throughout the experiment at a randomized order.

The experimental apparatus for the study used an individual testing room for each participant with a Mac computer, a PC keyboard, and a wired mouse. For the experiment, the lights in the room were turned off. A MATLAB (version x.) computer program designed by the lab was used for training, stimulus, and recording responses.

Procedure

Training

Before beginning the experiment, participants were instructed on the task through a script following the signing of the informed consent document. Once the participants indicated they understood what task they would be completing, the participants were led into the testing booth and completed the training section.

The training section presented participants with a multicolor cloud of dots. Each dot was colored red, green, or blue. One dot color was dominant, as defined as the color having more dots in the cloud than the other colors. The participant in the study indicated which color was the dominant color. Afterwards, the participant would indicate their level of confidence using a 1-4 scale, where 1 is the lowest confidence and 4 is the highest.

For the first three sections of training, the participants received feedback after the confidence rating on whether the color choice was correct or incorrect. For the first section, the cloud stimulus was presented for 3 seconds. Each training section shortened the duration of the stimulus, such that the duration of the second section was 1.5 seconds, and the third section was 0.5 seconds. After the first section, participants were reminded to fix their eyes on the white fixation dot in the center of the screen. After the second section, participants were reminded to use the whole scale for confidence ratings. For the fourth training section, participants practiced without feedback with stimulus presented for 0.5 seconds.

Experiment

Participants then completed the experimental section following the same procedures from the fourth training section. Participants completed 3 runs per session with 5 blocks per run. Each block had 48 trials. Each participant completed 720 trials per session. There was a 15-second break between each block, and participants chose how much time to take for a break between runs. This process took about an hour to complete and was completed twice for two sessions per participant. The second session was completed on another day within a week of the first session and included both a training and experimental section. Each participant completed 1440 experimental trials overall. Over 30 participants, this resulted in 21600 experimental trials, with 1800 trials for each of the 12 conditions.

Data Analysis

If the amount of overall information strength presented in a stimulus has a direct effect on confidence, then presenting a high and low level of overall information strength to participants will result in corresponding high and low confidence ratings. However, if the differences between information strength for each of the three stimuli choices also have a direct effect on confidence, then presenting higher and lower differences of information strength between the correct choice and the incorrect choices will result in corresponding higher and lower confidence ratings.

We tested four statistical hypotheses in this study involving stimulus strength and confidence ratings. The first hypothesis was that increased overall strength of the stimulus would cause increased confidence ratings. The second hypothesis was that increased difference between dominant color strength and secondary color strength would cause increased confidence ratings. The third hypothesis was that increased difference between dominant color strength and tertiary color strength would cause increased confidence ratings. The fourth hypothesis was that the overall strength, the difference between dominant and secondary color strength, and the difference between dominant and tertiary color strength would interact to modulate the confidence ratings so that higher levels of each variable would cause higher confidence ratings.

In order to test the four hypotheses, we conducted a 2 x 2 x 3 within-group repeated measures analysis of variance (ANOVA) because we had three categorical independent variables and one continuous dependent variable. The independent variables were the overall strength of stimulus (two levels: low or high), the difference between dominant and secondary color strength (two levels: low or high), and the difference between dominant and tertiary color strength (three levels: low, medium, or high). The dependent variable was the confidence rating on a Likert scale of 1 to 4. This 3-Way ANOVA allowed us to analyze the main effect of overall stimulus strength on confidence (Hypothesis 1), of the difference between dominant and secondary color strength on confidence (Hypothesis 2), of the difference between dominant and tertiary color strength on confidence (Hypothesis 3), and of the interactive effects of overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference on confidence (Hypothesis 4).

Results

Descriptives

The mean confidence rating across participants (*M* = 2.5695, *SD* = 0.9926) was normally distributed, suggesting there were no outliers in confidence scores. Using a two-sample t-test to determine gender differences, we found that the confidence ratings of female participants (*M* = 2.4093, *SD* = 0.1231) were not significantly different compared to that of male participants (*M* = 2.4228, *SD* = 0.0839), $t(13)=-0.2450$, $p = 0.8103$. Using another two-sample t-test to determine vision differences, we found that the confidence ratings of normal vision participants (*M* = 2.4029, *SD* = 0.0840) were not significant different compared to that of corrected-to-normal vision participants (*M* = 2.4504, *SD* = 0.1547), $t(13)=-0.7769$, $p = 0.4511$. Using a one-way ANOVA, there were no statistically significant difference between confidence ratings of the three colors red (*M*= 2.4311, *SD* = 0.1342), green (*M*= 2.4409, *SD* = 0.0820), and blue (*M*= 2.3749, *SD* = 0.1052), $F(2,42) = 1.5960$, $p = 0.2147$. Using another one-way ANOVA, there was a statistically significant difference between accuracy rates for the three colors red (*M*= 0.7429, *SD* = 0.0262), green (*M*= 0.7944, *SD* = 0.0277), and blue (*M*= 0.6430, *SD* = 0.0402), $F(2,42) = 86.95$, $p = 0.000$.

Hypothesis Testing

Four statistical hypotheses were tested in this study to determine the effects of stimulus strength on confidence ratings in the perceptual decision-making task. With the ANOVA ran to determine effect size and significance, Mauchly's Test of Sphericity indicated that the assumption of sphericity had been violated, $\chi2(65) = 231.54$, $p < .0005$ and $\chi2(65) = 248.71$, $p < .0005$; therefore, a Greenhouse-Geisser correction was used on each of the hypotheses.

Hypothesis 1

The first hypothesis was that the increased overall strength of the stimulus would cause increased confidence ratings. There was a significant main effect of the level of the overall strength of the stimulus on the confidence rating, $F(1,14) = 3133.6$, $p = 0.00$, $\eta p2 = 0.996$. Trials with higher overall stimulus strength (*M* = 2.6425, *SD* = 0.9883) resulted in higher confidence ratings compared to trials with lower overall stimulus strength (*M* = 2.4965, *SD* = 0.9916).

There was a significant main effect of the level of the overall strength of the stimulus on the accuracy rates, $F(1,14) = 22.99$, $p = 0.00$, $\eta p2 = 0.621$. Trials with higher overall stimulus strength (*M* = 0.7271, *SD* = 0.4455) resulted in higher accuracy rates compared to trials with lower overall stimulus strength (*M* = 0.714, *SD* = 0.452).

Hypothesis 2

The second hypothesis was that the increased difference between dominant color strength and secondary color strength would cause increased confidence ratings. There was a significant main effect of the level of the dominant-secondary strength difference on the confidence rating, $F(1, 14) = 1087.8$, $p = 0.00$, $\eta p2 = 0.987$. Trials with higher difference between dominant and secondary color strengths ($M = 2.6147$, $SD = 0.9939$) resulted in higher confidence ratings compared to trials with lower difference in dominant-secondary color strengths ($M = 2.5243$, $SD = 0.9893$).

There was a significant main effect of the level of the dominant-secondary strength difference on the accuracy rates, $F(1, 14) = 6707$, $p = 0.00$, $\eta p2 = 0.998$. Trials with higher difference between dominant and secondary color strengths ($M = 0.788$, $SD = 0.409$) resulted in higher accuracy rates compared to trials with lower difference in dominant-secondary color strengths ($M = 0.654$, $SD = 0.476$).

Hypothesis 3

The third hypothesis was that the increased difference between dominant color strength and tertiary color strength would cause increased confidence ratings. There was a significant main effect of the level of the dominant-tertiary strength difference on the confidence rating, $F(2, 28) = 535.04$, $p = 0.00$, $\eta p2 = 0.975$. Trials with higher difference between dominant and tertiary color strengths ($M = 2.6858$, $SD = 0.9748$) resulted in higher confidence ratings compared to trials with lower difference in dominant-tertiary color strengths ($M = 2.4553$, $SD = 1.0027$) or medium difference in dominant-tertiary color strengths ($M = 2.5674$, $SD = 0.9869$).

There was a significant main effect of the level of the dominant-tertiary strength difference on the accuracy rate, $F(2, 28) = 298.9$, p = 0.00, $\eta p2 = 0.955$. Trials with higher difference between dominant and tertiary color strengths (M = 0.758, SD = 0.429) resulted in higher confidence ratings compared to trials with lower difference in dominant-tertiary color strengths (M = 0.672, SD = 0.470) or medium difference in dominant-tertiary color strengths (M = 0.733, SD = 0.443).

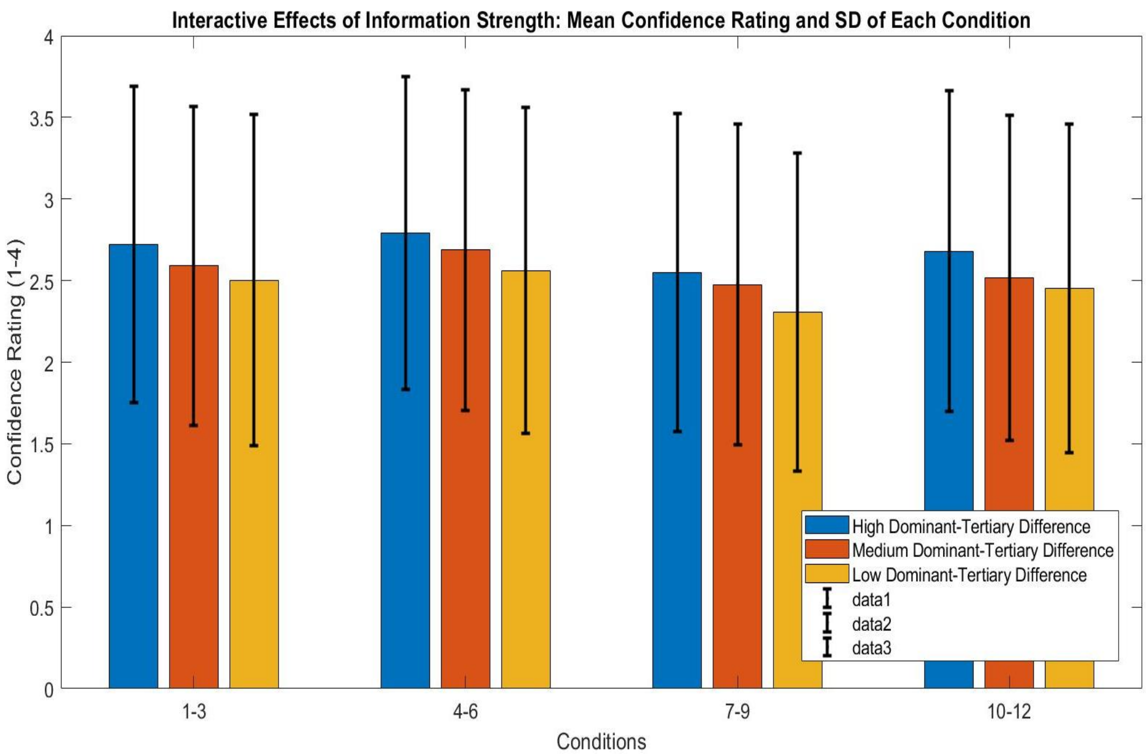
Hypothesis 4

The fourth hypothesis was that the overall strength, the difference between dominant and secondary color strength, and the difference between dominant and tertiary color strength would interact to modulate the confidence ratings so that higher levels of each variable would cause higher confidence ratings. There was a significant interaction between overall stimulus strength and dominant-secondary strength difference on the confidence in the decision, $F(1, 14) = 20.339$, $p = 0.0004$, $\eta p2 = 0.593$. There was a significant interaction between overall stimulus strength and dominant-tertiary strength difference, $F(2, 28) = 10.614$, $p = 0.0008$, $\eta p2 = 0.433$. There was also a significant interaction between dominant-secondary strength difference and dominant-tertiary strength difference, $F(2, 28) = 8.231$, $p = 0.0102$, $\eta p2 = 0.370$. Finally, there was a significant interaction between overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference together, $F(2, 28) = 11.751$, $p = 0.0025$, $\eta p2 = 0.457$. Varying levels of overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference did interact significantly with respect to confidence, as seen in Figure 2 below.

There was no significant interaction between overall stimulus strength and dominant-secondary strength difference on accuracy of the decision, $F(1, 14) = 4.0585$, $p = 0.0636$, $\eta p2 = 0.224$. There was a significant interaction between overall stimulus strength and dominant-tertiary strength difference, $F(2, 28) = 4.3719$, $p = 0.0449$, $\eta p2 = 0.235$. There was a significant interaction between dominant-secondary strength difference and dominant-tertiary strength difference, $F(2, 28) = 16.9136$, $p = 0.00$, $\eta p2 = 0.548$. Finally, there was no significant interaction between overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference together, $F(2, 28) = < 1$, p = 0.6901, $\eta p2 = 0.019$. Varying levels of overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference did interact significantly with respect to accuracy, as seen in Figure 3 below.

Figure 2.

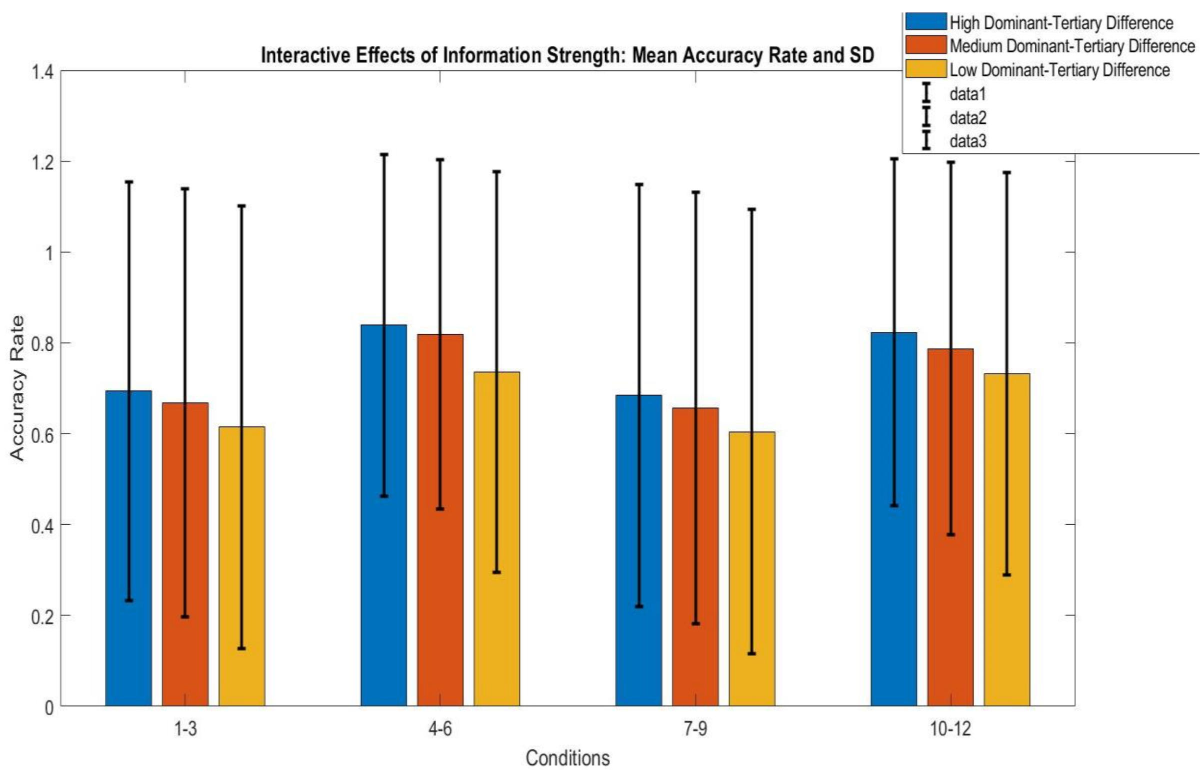
Interactive Effects of Information Strength: Mean Confidence Rating and Standard Deviations of Each Condition



Note. Conditions 1-12 are displayed in groups of three to compare differences in confidence for the three variables of overall strength, dominant-secondary strength, and dominant-tertiary strength. Each interaction between each variable was significant; this suggests that the main effects of each variable on confidence also interact with each other to create an interaction effect on confidence. This is demonstrated by condition 4 having a higher confidence than condition 1, condition 5, and condition 10.

Figure 3.

Interactive Effects of Information Strength: Mean Accuracy Rate and Standard Deviations of Each Condition



Note. Conditions 1-12 are displayed in groups of three to compare differences in accuracy for the three variables of overall strength, dominant-secondary strength, and dominant-tertiary strength. There were only significant interactions between overall strength and dominant-tertiary strength and more significantly between dominant-secondary strength and dominant-tertiary strength. This is demonstrated by a larger difference between condition 7 and condition 10 than condition 7 and condition 1.

Discussion

Hypotheses 1, 2, and 3 were supported by the data collected because significant main effects were found in the overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference and their positive correlation with confidence ratings. There was a significant positive correlation between overall stimulus strength and participants' confidence ratings. There was a significant positive correlation between dominant-secondary color strength differences and participants' confidence ratings. There was a significant positive correlation between dominant-tertiary color strength differences and participants' confidence ratings. Hypothesis 4 was supported by the data collected because there were significant positive interactive effects found in the overall stimulus strength, dominant-secondary strength difference, and dominant-tertiary strength difference on confidence ratings.

Theoretical Implications

Previous literature focuses on the effect of irrelevant information strength on the confidence and accuracy on tasks involving relevant perceptual information. In these studies, irrelevant stimuli had a positive relationship with confidence and no strong relationship with accuracy. However, these studies did not indicate the effects of relevant stimuli on confidence or accuracy. This study's results could be interpreted through two features in regard to information strength. First, overall stimulus strength had more effect on confidence than accuracy, so that more overall stimulus resulted in higher confidence. Second, the difference between the information strength for the two most prominent colors had a large effect on accuracy, so that a larger difference resulted in higher accuracy.

From this, two interpretations arise. Either the relevant stimuli have a positive relationship with accuracy or with confidence. If relevant stimuli are interpreted as the primary color through the difference between primary and secondary color information strength, then it results in the greatest change of accuracy compared to the other two independent variables. If relevant stimuli are interpreted as all stimuli through the overall stimulus strength, then it results in the greatest change of confidence compared to the other two independent variables. Understanding what the brain determines relevant or irrelevant in a multiple-alternative forced choice task will define the way computational models distinguish the confidence process. This also features in comparing the Bayesian holistic probability of events and Li and Ma's differential probability model (Zylberberg et al., 2014; Li & Ma, 2020). The results of this study indicate that both models may represent the data, but the effect sizes of the difference between primary and secondary colors on accuracy suggest that the differential probability model is more likely to represent confidence data in perceptual decision making tasks.

Strengths and Limitations

This study featured many developments in the field of perceptual decision making. It established a three alternative forced choice (3AFC) task unique to the field, and it demonstrated significant effects with strong effect sizes across conditions. The study features an experimental study with exploratory methods to develop the groundwork for future experiments and models.

However, the study is also limited in many features. One limitation is found in the descriptive statistics, where there was a significant difference in accuracy between blue and the other colors, which may have interacted with the results. In terms of wavelength, blue light is about 430nm, while red and green light are closer at about 570nm and 540nm, respectively; additionally, of the three photoreceptors within the retina, less than 10% are s-cones, or the neurons that process short wavelengths of light like the color blue (Calkins, 2001). These color values for red, green, and blue have all been used in perceptual research without any indication of significant differences; however, if previous designs did not compare only these three colors, and in terms of accuracy in a similar perceptual task, this distinction between blue and the other two colors may not have been established. Additionally, there was no significant difference in confidence between blue and the other colors. This could indicate that the metacognitive process of confidence overcompensates an individual's perspective on their ability to accurate perceive blue. In order to confirm this result is representative of the population's perceptual accuracy between the three primary colors of light, this aspect of the experiment should be replicated.

This study also did not progress to a computational model, but the design allows for future experiments to explore the implications of increments of variance, not just levels. Finally, the COVID-19 pandemic disrupted the research process, resulting in only 15 participants completing the full study and limited the time to establish a computational model within this specific study. This research creates an opportunity to investigate variance.

Future Research

Future research has many avenues to build off of this experiment. First, future studies might replicate this experiment with more participants to ensure a real representation of the population. Next, a probabilistic, normative computational model should be created, based off of posterior probability distributions, and then directly compared to the typical Bayesian confidence model and the Li and Ma difference model (Zylberberg et al., 2014; Li & Ma, 2020). In order to further investigate how the brain establishes stimuli as relevant or irrelevant in a perceptual decision making task, future studies should replicate this experiment with definitively irrelevant stimuli to compare to the overall stimulus strength and differential strengths in terms of effect on confidence and accuracy. Finally, more perceptual decision making studies should be conducted with 3AFC tasks, including previous studies in the literature.

Conclusion

The results of this study indicate that the difference between the two most prominent alternatives' information strengths is the most relevant factor for the metacognitive process of confidence in perceptual decision making. The overall strength and the difference between the primary and tertiary alternatives also have strong effects on confidence and should be factored in for future computational models. This may help in developing effective resources for situations with multiple choices to help workers choose the correct choice with the most confidence. Ultimately, this research on multiple alternatives in perceptual decision making tasks differs largely from the assumptions in the field about the confidence process, either in the statistical representations or in the assumptions on what constitutes relevant stimuli in an environment. Further investigation of these components must be conducted before the creation of computational models in order to establish the psychological process of confidence in applicable theory rather than idealized philosophy.

References

Boldt, A., de Gardelle, V., & Yeung, N. (2017). The impact of evidence reliability on sensitivity and bias in decision confidence. *Journal of Experimental Psychology: Human Perception and Performance, 43*(8), 1520-1531. doi:10.1037/xhp0000404.

Calkins, D. J. (2001). Seeing with S cones. *Progress in retinal and eye research*, *20*(3), 255-287. doi:10.1016/S1350-9462(00)00026-4

de Gardelle, V., & Mamassian, P. (2015). Weighting mean and variability during confidence judgments. *PloS one*, *10*(3). doi:10.1371/journal.pone.0120870

De Bondt, W. F., & Thaler, R. H. (1990). Do security analysts overreact?. *The American Economic Review*, 52-57.

Desender, K., Boldt, A., & Yeung, N. (2018). Subjective confidence predicts information seeking in decision making. *Psychological science*, *29*(5), 761-778. doi:10.1177/0956797617744771

Ferrigno, S., Kornell, N., & Cantlon, J. F. (2017). A metacognitive illusion in monkeys. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1862), 20171541. doi:10.1098/rspb.2017.1541

Kiani, R., Corthell, L., & Shadlen, M. N. (2014). Choice certainty is informed by both evidence and decision time. *Neuron*, *84*(6), 1329-1342. doi:10.1016/j.neuron.2014.12.015.

Koizumi, A., Maniscalco, B., & Lau, H. (2015). Does perceptual confidence facilitate cognitive control?.*Attention, Perception, & Psychophysics, 77*(4), 1295-1306. doi:10.3758/s13414-015-0843-3

Li, H. H., & Ma, W. J. (2020). Confidence reports in decision-making with multiple alternatives violate the Bayesian confidence hypothesis. *Nature Communications, 11*(1), 1-11. doi: 10.1038/s41467-020-15581-6

Meyniel, F., Sigman, M., & Mainen, Z. F. (2015). Confidence as Bayesian probability: From neural origins to behavior. *Neuron*, *88*(1), 78-92. doi: 10.1016/j.neuron.2015.09.039

Samaha, J., Barrett, J. J., Sheldon, A. D., LaRocque, J. J., & Postle, B. R. (2016). Dissociating perceptual confidence from discrimination accuracy reveals no influence of metacognitive awareness on working memory. *Frontiers in Psychology*, *7,* 851. doi:10.3389/fpsyg.2016.00851

Sanders, J. I., Hangya, B., & Kepecs, A. (2016). Signatures of a statistical computation in the human sense of confidence. *Neuron, 90*(3), 499-506. doi: 10.1016/j.neuron.2016.03.025

Spence, M. L., Dux, P. E., & Arnold, D. H. (2016). Computations underlying confidence in visual perception. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(5), 671. doi:10.1037/xhp0000179

Spence, M. L., Mattingley, J. B., & Dux, P. E. (2018). Uncertainty information that is irrelevant for report impacts confidence judgments. *Journal of experimental psychology: human perception and performance*, *44*(12), 1981. doi:10.1037/xhp0000584

Vlassova, A., Donkin, C., & Pearson, J. (2014). Unconscious information changes decision accuracy but not confidence. *Proceedings of the National Academy of Sciences*, *111*(45), 16214-16218. doi:10.1073/pnas.1403619111

Zylberberg, A., Fetsch, C. R., & Shadlen, M. N. (2016). The influence of evidence volatility on choice, reaction time and confidence in a perceptual decision. *Elife*, *5*, e17688. doi:10.7554/eLife.17688

Zylberberg, A., Roelfsema, P. R., & Sigman, M. (2014). Variance misperception explains illusions of confidence in simple perceptual decisions. *Consciousness and cognition*, *27*, 246-253. doi:10.1016/j.concog.2014.05.012